Available online at www.sciencedirect.com
SCIENCE DIRECT®
Applied Acoustics 66 (2005) 1262–1277
applied acoustics
www.elsevier.com/locate/apacoust

Least-error localization of discrete acoustic sources

David C. Torney [a,*], Robert J. Nemzek [b]

[a] *Theoretical Division, Mailstop K710, Los Alamos National Laboratory, Los Alamos, NM 87545, United States*
[b] *ISR Division, Mailstop D448, Los Alamos National Laboratory, Los Alamos, NM 87545, LAUR # 03-5797, United States*

Received 1 October 2003; received in revised form 5 January 2005; accepted 18 February 2005
Available online 2 June 2005

Abstract

We describe a new, least-error method for locating a discrete acoustic source (which generates a radially symmetric, outgoing wave) based on time-of-arrival data. This method localizes the source by minimizing the sum of the absolute values of the differences between the squares of the theoretical and actual times of arrival. The method is suited to noisy data, and whenever the errors in the data are unbiased, the more times of arrival used, the greater the expected accuracy of localization. The method is simplest for two dimensional data, requiring only elementary algebra. By means of simulations, we demonstrate the amelioration of localization with the number of times of arrival employed: the average inaccuracy falls asymptotically as the reciprocal of the square root of this number. The new method also yielded more accurate localization, on the average, than a least-square method. We make direct comparison with time-difference-of-arrival localizations, both for simulated data and for experimental data collected at a shooting range, demonstrating the favorability of the new method. We also demonstrate its facilitation of the localization of multiple, cotemporary sources: via partitioning of the data. Our method is suited to sensor networks with computationally empowered nodes.

* Corresponding author. Tel.: +1 505 667 7510; fax: +1 505 665 3493.
E-mail address: dct@lanl.gov (D.C. Torney).

0003-682X/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.apacoust.2005.02.010

Keywords: Distributed sensor networks; Lagrange multipliers; Multiple sources; Optimization; Quadratic formula; Time-difference-of-arrival; Unbiased estimator

1. Introduction

Distributed sensor networks (DSN) have computationally empowered nodes and may contain manifold sensors. DSNs have much potential for surveillance. An advantageous preliminary exercise is implementing them for "discrete acoustic-source localization", as ulterior aims are more challenging (cf. [13]).

Let a sensor network have S nodes at positions $\mathbf{r}_s$; $s = 1, 2, \ldots, S$, each node comprising an acoustic sensor. Let there be a discrete sound source at position $\mathbf{r}$ and time t, and, whence, an outgoing, radially symmetric, shock wave. Therefore, in the ideal, the time-of-arrival at node s, $t_s = t + \|\mathbf{r}_s - \mathbf{r}\|/v$; $s = 1, 2, \ldots, S$, where v denotes the velocity of sound, assumed constant, and where $\|\mathbf{w}\|$ denotes the Euclidean length of the vector $\mathbf{w}$.

The objective is to predict $\mathbf{r}$ and t based on the t_s's; thus, our formulation does not apply to more sophisticated data, such as those collected by sensors determining the source's orientation [10]. As v is taken to be constant, we subsequently eliminate t's in favor of $d \stackrel{\text{def}}{=} vt$ and $d_s \stackrel{\text{def}}{=} vt_s; \quad s = 1, 2, \ldots, S$. Therefore, for $d' > d$, the locus $\mathbf{r}'(d')$ of the wave is

$$d' - d = \|\mathbf{r}'(d') - \mathbf{r}\|. \tag{1}$$

In this ideal setting, enough d_s's should suffice for obtaining d and $\mathbf{r}$, but how many? $S = 1$ is essentially useless. In two spatial dimensions, geometrical considerations yield that the locus of the spatial coordinates of the source is a hyperbola for $S = 2$ (taking the difference of two (1)'s, for, say, $\mathbf{r}'$ and $\mathbf{r}''$), two points for $S = 3$ (viz. Section 2.2) and one point for $4 \leqslant S$. In fact, $S = 4$ suffices, in the absence of errors, to locate the source in (2-d) space and time; such a formulation is referred to as the time-difference-of-arrival (TDOA) method [1,11] (cf. Appendix A). This method is essential for GPS [7] and cell-phone localizations [15].

These elementary approaches are, however, plainly unsuitable for non-ideal data, comprising, for instance, measurement errors and, also, the influences of impediments. Mitigation is afforded by least-error methods, which yield a solution $\mathbf{r}$ and d by minimizing the sum, over all nodes, of "discrepancies" between the sides of (1) [7,8,18]. When systematic errors have been eliminated, the accuracy of least-error solutions should increase with S.

Herein we develop a new least-error method based on measuring discrepancies by absolute values. Minimizing the sum of the absolute values of the "errors" was advocated by Laplace [4, p. ix]. Minimizing the sum of these summands, instead of the sum of the squares of the errors, better discounts outlying measurements. For select objectives, such as the current one, it can be simpler, mathematically, to optimize the sum of the absolute values of the errors – despite engendering discontinuous derivatives [17]. Note that a natural least-square method requires finding the roots of a polynomial of the sixth degree [8]. As described in Section 2, in two spatial dimensions, our least-error method requires only elementary algebraic methods – comparable to those used by the TDOA method [11] (viz. Appendix A). The increased complexity required for implementing the method three spatial dimensions is broached in Appendix B.

We apply our least-error methodology in Section 3. There, simulations demonstrate the amelioration of localization afforded by increasing the number of nodes; its localization appears to be unbiased and to be more accurate than a comparable least-square method [18] (viz. Section 3.1). We also compare its performance with that of the TDOA method: using four nodes in a plane, both with simulated data and with experimental data. We also demonstrate the facilitation, by our method, of the localization of multiple, cotemporary discrete sources.

2. Source localization in two spatial dimensions

In this section we describe our methods, specialized to localization within a plane. Let there be given S nodes with their measured positions denoted by $\mathbf{r}_s = (x_s, y_s)$; $s = 1, 2, \ldots, S$. The source is assumed to emit a cylindrical wave – emanating from $\mathbf{r} = (x, y)$ "at" d. The data recorded at the nodes are the respective d_s's; $s = 1, 2, \ldots, S$; from these one determines $\mathbf{r}$ and d.

Our measure of the "discordance" of the data, given x, y and d, is

$$\Delta(x, y; d) \stackrel{\text{def}}{=} \sum_{s=1}^{S} w_s |(x - x_s)^2 + (y - y_s)^2 - (d - d_s)^2|, \tag{2}$$

where the w_s's, introduced for sake of generality, are positive, real constants. The main motivation for (2) is that it is easy to optimize. Candidate location(s) of the source are x, y and d at the global minimum (minima) of Δ.

As the optimization of Δ is unconstrained, it is necessary for the first partial derivatives of Δ to vanish at its global optimum – provided these derivatives exist. Recall that the derivative of $|f(x)|$ equals $f'(x)\,\mathrm{sgn}\,(f(x))$, with $\mathrm{sgn}\,(y)$ denoting the signum function (taking values $-1, 0$ and 1 depending on whether $y \in \mathbb{R}$ is less than, equal to or greater than zero, respectively). This derivative exists when $f(x) \neq 0$ or $f'(x) = 0$. Therefore, at a stationary point of (2):

$$\frac{\partial \Delta}{\partial x} = 2 \sum_{s=1}^{S} w_s (x - x_s) \mathrm{sgn}_s = 0;$$

$$\frac{\partial \Delta}{\partial y} = 2 \sum_{s=1}^{S} w_s (y - y_s) \mathrm{sgn}_s = 0;$$

$$\frac{\partial \Delta}{\partial d} = -2 \sum_{s=1}^{S} w_s (d - d_s) \mathrm{sgn}_s = 0,$$

where sgn_s denotes $\mathrm{sgn}\,((x - x_s)^2 + (y - y_s)^2 - (d - d_s)^2)$; $s = 1, 2, \ldots, S$.

These partial derivatives are not defined at points where any sgn_s vanishes. When no sgn_s vanishes, consideration of the second partial derivatives establishes that any stationary point obtained by solving the foregoing system is a saddle point: never a minimum. To find the minima of Δ one must, therefore, confront some subtleties of optimizing non-differentiable functions.

Upon reflection, minima of Δ are found to be restricted to manifolds in x, y, d-space where one or more sgn_s vanish (the ones mentioned in Section 1). With such restrictions, the foregoing derivatives are well defined (at loci where no additional summands vanish), as one may clearly omit summands which vanish everywhere. It is immediate that at least two sgn_s's must vanish at a local minimum of Δ. Cases with two vanishing sgn_s's and three vanishing sgn_s's are treated separately.

2.1. Two sgn_s's zeroed (2-d)

Here we exhibit all the stationary points of Δ resulting when pairs of sgn_s's are zeroed. The global minimum must occur either at such points or at the stationary points resulting from zeroing triples of sgn_s's, described in Section 2.2.

Two Lagrange multipliers – λ and μ – facilitate finding the stationary points resulting from zeroing two sgn_s's: sgn_ℓ and sgn_m, say, with $\ell \neq m$. Therefore, consider the stationary points of

$$\tilde{\Delta} = \Delta + \lambda((x - x_\ell)^2 + (y - y_\ell)^2 - (d - d_\ell)^2) + \mu((x - x_m)^2 + (y - y_m)^2 - (d - d_m)^2). \tag{3}$$

Here, ℓ is not equal to m, and, to obtain this class of stationary points, $\{\ell, m\}$ ranges over all unordered pairs of indices from $\{1, 2, \ldots, S\}$. Thus, given none of the remaining sgn_s's vanish, the following first partial derivatives exist:

$$\frac{\partial \tilde{\Delta}}{\partial x} = 2\sum_{s=1}^{S}(x - x_s)w_s\mathrm{sgn}_s + 2\lambda(x - x_\ell) + 2\mu(x - x_m) = 0;$$

$$\frac{\partial \tilde{\Delta}}{\partial y} = 2\sum_{s=1}^{S}(y - y_s)w_s\mathrm{sgn}_s + 2\lambda(y - y_\ell) + 2\mu(y - y_m) = 0;$$

$$\frac{\partial \tilde{\Delta}}{\partial d} = -2\sum_{s=1}^{S}(d - d_s)w_s\mathrm{sgn}_s - 2\lambda(d - d_\ell) - 2\mu(d - d_m) = 0.$$

This system is composed of one linear equation in each unknown: x, y and d. λ and μ will subsequently be selected to implement the two constraints $\mathrm{sgn}_\ell = \mathrm{sgn}_m = 0$. If solution is feasible, then, this solution is a stationary point of Δ.

The foregoing three equations yield, respectively,

$$x = \frac{\xi_x + \lambda x_\ell + \mu x_m}{\xi + \lambda + \mu}, \quad y = \frac{\xi_y + \lambda y_\ell + \mu y_m}{\xi + \lambda + \mu} \quad \text{and } d = \frac{\xi_d + \lambda d_\ell + \mu d_m}{\xi + \lambda + \mu},$$

denoting $\sum_{s=1}^{S} w_s\mathrm{sgn}_s$ by ξ, $\sum_{s=1}^{S} x_s w_s\mathrm{sgn}_s$ by ξ_x, $\sum_{s=1}^{S} y_s w_s\mathrm{sgn}_s$ by ξ_y and $\sum_{s=1}^{S} d_s w_s$-sgn_s by ξ_d. Note that although the sgn_s's are functions of the unknowns, one may proceed by substituting each possible $(-1, +1)$-binary string of length $S - 2$ and seeking self-consistency (substituting a solution, x, y and d, into the sgn_s's; $s = 1, 2, \ldots, S$). Furthermore, the constraints yield two quadratic equations: one in λ and the other in μ:

$$(\xi_x - \xi x_m + \lambda(x_\ell - x_m))^2 + (\xi_y - \xi y_m + \lambda(y_\ell - y_m))^2 \\ = (\xi_d - \xi d_m + \lambda(d_\ell - d_m))^2; \tag{4}$$

$$(\xi_x - \xi x_\ell + \mu(x_m - x_\ell))^2 + (\xi_y - \xi y_\ell + \mu(y_m - y_\ell))^2 \\ = (\xi_d - \xi d_\ell + \mu(d_m - d_\ell))^2. \tag{5}$$

It is easily seen that the roots of (4) and (5) share some key attributes:

$$\lambda_\pm = (-b_\lambda \pm \sqrt{D})/2a, \tag{6}$$

and

$$\mu_\pm = (-b_\mu \pm \sqrt{D})/2a, \tag{7}$$

where D denotes the discriminant common to these two equations, with

$$D/4 = ((x_\ell - x_m)(\xi_d - \xi d_m) - (d_\ell - d_m)(\xi_x - \xi x_m))^2 + ((y_\ell - y_m)(\xi_d - \xi d_m) \\ - (d_\ell - d_m)(\xi_y - \xi y_m))^2 - ((x_\ell - x_m)(\xi_y - \xi y_m) - (y_\ell - y_m)(\xi_x - \xi x_m))^2,$$

where $a = (x_\ell - x_m)^2 + (y_\ell - y_m)^2 - (d_\ell - d_m)^2$ denotes the coefficient of the quadratic term, also common to the two equations; and where b_λ and b_μ denote the respective coefficients of the linear terms. Whenever $D < 0$, the Lagrange multipliers are complex and the corresponding solutions cannot be stationary points [16, Theorem 2.6].

For a stationary point (x, y, d) to be admissible, its components must all be finite. When the denominator $\xi + \lambda + \mu$ vanishes, in all likelihood, some components are infinite. Consider, for example, the linear system engendered under the assumption that all are finite:

$$\begin{aligned} \xi_x + \lambda x_\ell + \mu x_m &= 0, \\ \xi_y + \lambda y_\ell + \mu y_m &= 0, \\ \xi_d + \lambda d_\ell + \mu d_m &= 0. \end{aligned} \tag{8}$$

According to the fundamental theorem on overdetermined linear systems [6, vol. 7, p. 52], this system will admit a solution (λ, μ) if and only if

$$\text{Rank}\begin{pmatrix} x_\ell & x_m \\ y_\ell & y_m \\ d_\ell & d_m \end{pmatrix} = \text{Rank}\begin{pmatrix} x_\ell & x_m & \xi_x \\ y_\ell & y_m & \xi_y \\ d_\ell & d_m & \xi_d \end{pmatrix}.$$

Given the random measurement errors comprised, the rank of the left-hand matrix will almost always be smaller than the rank of the right-hand matrix. Assuming these ranks are unequal, the system (8) has no solution. Therefore, when the denominator $\xi + \lambda + \mu$ vanishes, at least one of x, y and d will almost surely be infinite, and such solutions may be ignored because they cannot yield a minimum value of Δ.

If additional sgn_s's were to vanish at one of the foregoing stationary points, then such points are not admitted to this class.

2.2. Three sgn_s's zeroed (2-d)

Note that, with noisy data, it should almost never be the case that more than three sgn_s's jointly vanish, and it is safe to ignore this potentiality. When zeroing three sgn_s's, the use of Lagrange multipliers leads to a system whose solution is cumbersome, as corroborated by Appendix B, and, therefore, direct elimination is preferable. As is the case with the stationary points resulting from zeroing pairs of sgn_s's, zeroing each triple will be seen to yield either zero, one or two stationary points of Δ.

Consider the following quadratic system:

$$(x - x_\ell)^2 + (y - y_\ell)^2 = (d - d_\ell)^2; \tag{9}$$

$$(x - x_m)^2 + (y - y_m)^2 = (d - d_m)^2; \tag{10}$$

$$(x - x_n)^2 + (y - y_n)^2 = (d - d_n)^2, \tag{11}$$

where, ℓ, m and n are distinct elements of $\{1, 2, \ldots, S\}$. Taking the differences (9), (10) and (9)–(11) yields two linear equations in x, y and d. This allows, say, y and d to be expressed in the following forms:

$$y = \alpha x + \beta; \tag{12}$$

$$d = \gamma x + \delta, \tag{13}$$

where

$$\alpha = \frac{\frac{x_\ell - x_m}{d_\ell - d_m} - \frac{x_\ell - x_n}{d_\ell - d_n}}{\frac{y_\ell - y_n}{d_\ell - d_n} - \frac{y_\ell - y_m}{d_\ell - d_m}},$$
$$\beta = \frac{\psi_m - \psi_n}{\frac{y_\ell - y_n}{d_\ell - d_n} - \frac{y_\ell - y_m}{d_\ell - d_m}},$$
$$\gamma = \frac{x_\ell - x_m}{d_\ell - d_m} + \alpha \frac{y_\ell - y_m}{d_\ell - d_m},$$
$$\delta = \beta \frac{y_\ell - y_m}{d_\ell - d_m} + \psi_m,$$

with

$$\psi_j = \frac{x_j^2 - x_\ell^2 + y_j^2 - y_\ell^2 + d_\ell^2 - d_j^2}{2(d_\ell - d_j)}; \; j \in \{m, n\}.$$

Safeguards against the vanishing (or near vanishing) of the denominators in the expressions for the foregoing coefficients should be given consideration. Under appropriate circumstances, one might want to implement the computation of the "limiting behavior", or, if necessary, extended floating-point arithmetic. For our applications, described in Section 3, double precision sufficed.

Substituting (12) and (13) into (9) yields a quadratic equation in x:

$$ax^2 + bx + c = 0, \tag{14}$$

with

$$a = 1 + \alpha^2 - \gamma^2,$$
$$b = -2(x_\ell - \alpha(\beta - y_\ell) + \gamma(\delta - d_\ell)) \text{ and}$$
$$c = x_\ell^2 + (\beta - y_\ell)^2 - (\delta - d_\ell)^2.$$

Solving (14) for x yields, when it is real, one or two stationary points (x, y, d).

2.3. Global minimization of Δ

Global minimization of Δ involves all pairs and triples of nodes: zeroing their sgn_s's. Respective stationary points are generated, as described above, and the minimum value of Δ over these stationary points is its global minimum.

For pairs of zeroed sgn_s's, one ranges over $\binom{S}{2}$ pairs of indices. For each pair, one must also range over the 2^{S-2} possibilities for the nonzero sgn_s's (to postulate the ξ's). For each pair and each choice, candidate stationary points are generated. These may be admitted if: (i) their λ and $\mu \in \mathbb{R}$, (ii) their sgn_s's reproduce the postulated sgn_s's and (iii) no additional sgn_s's vanish. Thus, the computation of these stationary points involves the data from all nodes.

For triples of zeroed sgn_s's, the stationary-point generation procedure is more straightforward. One must range over the $\binom{S}{3}$ triples of indices. For each triple, one solves (14) and, the real roots yielding admissible stationary points without reference to the data from the remaining nodes; those data allow the ranking of the stationary points.

3. 2-d Applications of least-error methodology

In this Section, the w_s's are, for simplicity, taken equal to unity. In all of our applications, one x, y and d minimized Δ.

3.1. Localization errors as a function of S: comparison with LS

Consider a square of unit side with vertices $(0, 0)$, $(0, 1)$, $(1, 0)$ and $(1, 1)$ and a source at its center, $(1/2, 1/2)$, emitting a cylindrical wave at $d = 0$. S nodes were independently placed, uniformly and randomly, within the square. A random error, uniform on $[-.01, .01]$, was then added to the exact value of d_s, independently; $s = 1, 2, \ldots, S$.

First, x, y and d were obtained by application of the least-error method, described in Section 2.1. Second, we obtained these by application of least-square techniques based upon [18, Appendix B].[1]

[1] For every pair of nodes, indexed by $\{\ell, m\}$, with $\ell \neq m$, one takes the difference of respective equations obtained from (1): $(x_\ell - x)^2 + (y_\ell - y)^2 = (d_\ell - d)^2$, and $(x_m - x)^2 + (y_m - y)^2 = (d_m - d)^2$. This yields $\binom{S}{2}$ linear equations, whose least-square solution follows from their *normal equations*: three linear equations in x, y and d [6, vol. 5, pp. 376–380]. (As with the least-error method, all weights equalled unity.).

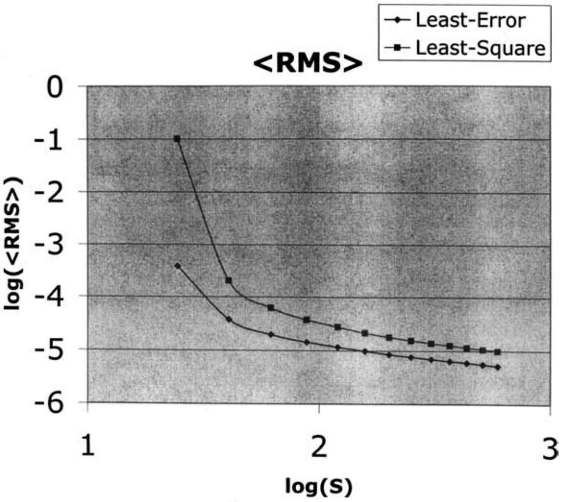

Fig. 1. $\langle \text{RMS} \rangle$ as a function of S; $4 \leqslant S \leqslant 16$. Abscissa: $1 \leqslant \log(S) \leqslant 3$; Ordinate: $\log(\langle \text{RMS} \rangle)$. Results for the least-error method are marked with diamonds; results for the least-square method are marked with squares. In both cases, the asymptotic slope is $-1/2$.

For both methods, we determined the respective average root-mean-square error:

$$\left\langle \sqrt{(x-1/2)^2 + (y-1/2)^2 + d^2} \right\rangle; \; 4 \leqslant S \leqslant 16.$$

(Uniqueness of the global minimum would be unlikely for $S < 4$.) These averages are plotted on logarithmic axes in Fig. 1.

Note that the average root-mean-square errors for both behaved asymptotically as $S^{-1/2}$, as these are the asymptotic slopes of the curves in Fig. 1. These results are rationalized in Section 4. The magnitude of the average root-mean-square errors also varied directly with the magnitude of the errors in the data.

3.2. Comparison with TDOA: simulation

In two dimensions, the TDOA method requires $S = 4$ [3,7,11,18] (viz. Appendix A). Note that this approach always yields a solution, but that it is ill adapted to noise.

We compared the accuracies of TDOA solutions and the least-error method, applied to a data set simulated as described in Section 3.1. Table 1 contains some error "percentiles" for the two computations. For example, for the TDOA method, 75% of the time, the root-mean-square (RMS) error, $\sqrt{\|\mathbf{r} - \mathbf{r}_0\|^2 + (d - d_0)^2}$, was less

Table 1
Error percentiles for root-mean-square localization ($S = 4$)

	$\langle \text{RMS} \rangle$	25%	50%	75%	95%
TDOA	–	0.008	0.014	0.035	0.218
Least-error	0.033	0.007	0.0099	0.0175	0.075

than 0.035, whereas the corresponding value for the least-error method was 0.0175. (The nought-subscripted indeterminates in this formula denote the "true values" and the non-subscripted indeterminates the solution values.) Due to a "heavy tail" in the TDOA results, its average RMS error, likely unbounded, could not be reliably estimated, even in a large number of trials ($>10^6$).

3.3. Comparison with TDOA: experiment

Prior to developing our least-error methods, we implemented and tested a prototype TDOA acoustic-source localization network. For this test, six acoustic-sensor nodes were arranged in two groups of three (whose members' locations are depicted as +'s in Figs. 2(a) and (b)). The system was fielded at the Los Alamos National Laboratory & Protective Technology Los Alamos firearms training facility. Each node contained an acoustic sensor, a microprocessor, an inexpensive GPS receiver and an RF transceiver. They jointly constituted a sensor network which autonomously located gunshots using the TDOA method.

The on-board GPS receivers provided both a common timebase for the network and the locations of the individual nodes. Inaccuracies in these positions, whose RMS error was roughly 10 m and which was strongly correlated across nodes, were the dominant errors in this setting. The time-of-arrival determination involved errors of ±0.1 ms; multiplying by v establishes that these errors ($\sim$0.03 m) were much smaller than the GPS errors, and additional errors were of comparably negligible magnitude.[2] For instance, the latency of the acoustic sensors had a standard deviation of approximately 0.05 ms, etc.

Each node performed a threshold detection of sound amplitude, recording the arrival time of the cylindrical wavefront. The arrival times were then propagated around the network using a flooding-style communications protocol. The first node to receive four times of arrival (typically its own measurement and three others) calculated the sound source position, using the TDOA algorithm detailed in Appendix A [11]. The implemented communications algorithm typically suppressed the transmission of the fourth time-of-arrival measurement: the one belonging to the node which performed the TDOA calculation, using its own measurement and that from three other nodes. Thus, to compare the TDOA method to our least-error method, we first identified TDOA localizations which could unambiguously be matched to a set of three measurements; then, from the three, the fourth measurement was reconstructed, using the uniqueness of the TDOA solution. These sets of four unambiguous measurements were analyzed using both methods.

Fig. 2(a) depicts the results of the TDOA method. Two shots were located within 3 m of their true position by the TDOA method, but this approach predicted that the majority of the shots occurred 10–30 m south of their true locations. As mentioned,

[2] v was taken to equal 343 m/s for these experiments.

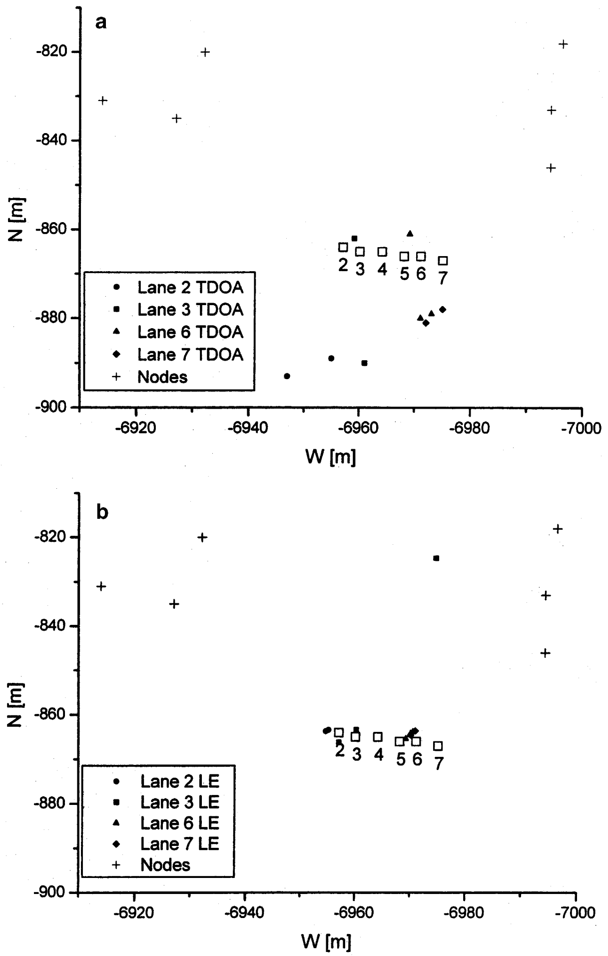

Fig. 2. (a) TDOA localizations using six nodes (+'s) and, for each gunshot, one quad of times of arrival. Lanes (numbered squares) correspond to shooters; the lanes used were #'s 2, 3, 6 and 7. The predicted coordinates for the gunshots for firings from each of these lanes are indicated by respective symbols. The dimensions of the axes are in meters, and the x-axis is oriented east–west, whereas the y-axis is oriented north–south; these are referenced to an arbitrary origin northwest of the plot. Note that the point from lane three at $(-6959, -862)$ is actually the superposition of two nearly identical predictions for coordinates of two gunshots from lane three. (b) Least-error localizations, using the same nodes and combinations of four times of arrival as in (a). The coordinate axes are also the same as in (a).

discrepancies derive from GPS position error, variation in individual acoustic threshold settings, dispersion of sound, etc.

Fig. 2(b) depicts the results of our least-error method. Almost all (eight of nine) of the firings were located within 3 m of their true position by our least-error method.

Table 2
The proportions of false positives in the below-threshold results for different numbers of nodes. False positives were defined as those with the minimum, over all sources $(\mathbf{r}_0; d_0)$: $\|\mathbf{r} - \mathbf{r}_0\|^2 + (d - d_0)^2 > 0.02$

S	False positive rate
4	0.4
6	0.2
8	0.03

3.4. Deconvolution of multiple sources

Least-error source localization may be also be used for the separation, or deconvolution, of multiple sources. Recall that with multiple, sinusoidally varying, plane-wave sources, established methods yield the parameter estimates for all sources based upon measurements at the nodes [9]. To locate multiple cylindrical-wave "point sources", based on times of arrival, constitutes a distinct challenge.

Were the times of arrival for multiple discrete sources received at the nodes, one could proceed by perform the least-error analysis for each combination: using one signal from each node. Thus, if there were n_s signals detected at the sth node, then the corresponding number of least-error calculations would be $\prod_s^S n_s$. The minimum values of Δ obtained should discriminate between true and artifactual sources, as it is expected that invalid combinations should exhibit large discordances.[3]

To illustrate these considerations, we modified the simulations of Section 3.1 to include three sources: at $d_0 = 0$ and at $(1/4, 1/4)$, $(1/2, 1/2)$ and $(3/4, 3/4)$; $n_s = 3$; $s = 1, 2, 3$. Noise was added as above. Only those combinations of signals with minimum $\Delta < 0.03$ were retained. (For other situations, alternative thresholds would plainly be appropriate; this choice pertains to the noise level in our simulations $\mathcal{O}(0.01)$.) The results of our simulations, involving $S = 4, 6$ and 8, are as follows; see Table 2.

For virtually all configurations of the nodes, all three sources were detected (and from multiple signal combinations). Note that having sources emit at substantially different d's would decrease the false positive rates. In Fig. 3, the spatial coordinates of the below-threshold results for $S = 8$ are depicted.

4. Discussion

The results of Section 3 illustrate various advantages of our least-error method over other methods. For example, it accrued better average performance than a least-square method, depicted in Fig. 1. Table 1 illustrates that our least-error method has higher accuracy than TDOA, when the data contain noise. Figs. 2(a) and (b) demonstrate its practical merits.

[3] This approach assumes most nodes records a time-of-arrival from each source. When this is not the case, one could attempt omission of the nodes receiving the smallest numbers of times of arrival.

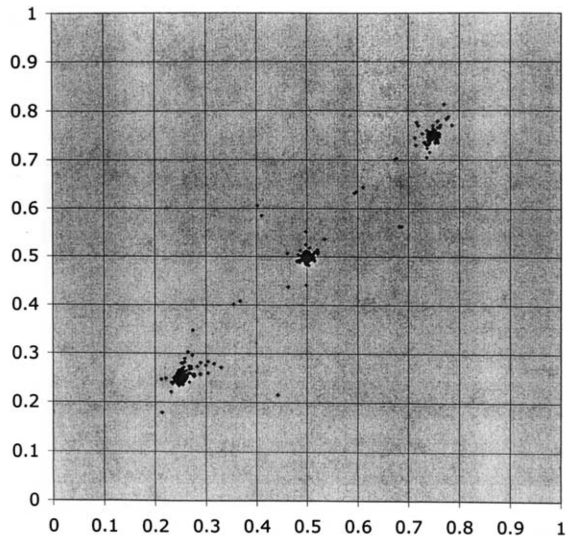

Fig. 3. The predicted spatial coordinates of the below-threshold (minimum $\Delta < 0.03$) least-error solutions, given three simultaneous sources, at $d = 0$, and (1/4, 1/4), (1/2, 1/2) and (3/4, 3/4).

A novel characteristic of our least-error method is that some of the discrepancies (sgn_s's) must vanish at the minimum. When fewer than the maximum number of sgn_s's vanish (these being three and four in two and three dimensions, respectively) our implementation of the method has computational complexity exponential in the number of nodes. It remains a challenge to attempt to circumvent the latter complexity. On the other hand, for the maximum numbers of vanishing sgn_s's, ours is evidently a polynomial algorithm which could readily be implemented over the nodes of a DSN. For some applications and for (yet to be determined) classes of noise, simply ranging over the latter stationary points and accepting the minimum value of Δ may afford suitable source localization. The average performance for the simulations of this manuscript would not noticeably change if the stationary points involving two zeroed sgn_s's were omitted.

By analogy to our DSN implementation of the TDOA method (cf. Section 3.3), our least-error approaches for acoustic-source localization in two spatial dimensions could also be implemented. Efficient communication protocols would need to be devised. The optimal value of Δ provides an indication of the accuracy of a localization, a basis for deciding whether to include the data from more nodes – or to report the current localization. Such an implementation, coupled with the ability of nodes to emit sound, would provide an alternative method for node localization.

For accessible S, the average RMS localization error of our method falls, asymptotically, as $S^{-1/2}$, as illustrated in Fig. 1. This reflects the optimal spatial resolution S sensors may afford. In detail, independent noise components contribute $\mathcal{O}(S^{1/2})$ to Δ, but alterations to the unknowns yield (coherent) contributions $\mathcal{O}(S)$. Therefore, the scale of "spatial resolution" of the network cannot exceed their quotient:

$\mathcal{O}(S^{-1/2})$. It is no surprise that least-error methods achieve this limit because their optima should fall within this "noise-limited" domain.

Due to dispersion, one might expect that an important component of the noise in the d_s's will increase with the Euclidean distance from the source. Using the w_s's for mitigation seems desirable, but it would greatly complicate the optimization of Δ because it would make the w_s's functions of the indeterminates. For many purposes it may suffice to iterate: solving a sequence of optimizations – each iterate having constant w_s's – based on the inferred $\mathbf{r}$ and d from the previous iteration. If, for example, the standard deviation σ_ℓ of $|(x-x_\ell)^2+(y-y_\ell)^2-(d-d_\ell)^2|$equals a known function of $\|\mathbf{r}-\mathbf{r}_0\|$, then one might want to take $w_s = 1/\sigma_s$; $s = 1, 2, \ldots, S$. Convergence of this procedure is not necessarily at issue, as one iteration could accrue greater accuracy than the use of equal w_s's.

More sophisticated methods are plainly required for localization of a large number of cotemporary sources. "Time-slicing", with overlapping time intervals, is a practical approach for reducing the n_s's. The slice length would engender "neighborhoods" of nodes which could collaborate in source localization. One might, furthermore, effect the desired importance sampling of the signal combinations by means of the Markov chain Monte Carlo method: transitioning between combinations in accordance with the likelihood ratio for the data. A useful approximation yielding the likelihood of the data could involve the inferred source coordinates. Such like should be put to the test.

When there are two real roots of (6) and (7), for $S = 2$, or of (14), for $S = 3$, these solutions may not, in general, be distinguished, even though only one solution could pertain to the source. Therefore, source localization requires either a different approach or larger values of S. This could, for some, inspire a higher regard for the sophistication of many creatures, such as bats and dolphins, endowed with only two acoustic sensors – but whose vocations (or avocations) employ echolocation [2,5,14,19].

Acknowledgements

This work was performed with NNSA and Los Alamos National Laboratory Discretionary Research support for Modelling and Testing of Distributed Sensor Networks with Collective Computation, Jared Dreicer, P.I., under the aegis of the University of California, under contract W-7405-ENG-36. We are indebted to Angela Mielke, who has replaced Jared Dreicer as P.I., for reading a draft of this manuscript. This manuscript is dedicated to the memories of our Town-School classmates and comrades Michael Coleman, David Geddes, Mark Gordon and Derrick Parmenter.

Appendix A. TDOA

Our definition of TDOA is derived from [11], which we reproduce for the convenience of the reader, using the present notation. It originates from the system of four equations derived by squaring (1) for $S = 4$. Take the difference of one of these with the other three, to obtain three linear equations in x, y and d. Eliminating d yields two linear equations in x and y, and solve these. d is then obtained by substitution of x and y. Here, d_{ij} denotes $d_i - d_j$, x_{ij} denotes $x_i - x_j$ and y_{ij} denotes $y_i - y_j$. In this way we obtain

$$x = \frac{B_2C_1 - B_1C_2}{A_1B_2 - A_2B_1}; \quad y = \frac{A_1C_2 - A_2C_1}{A_1B_2 - A_2B_1};$$
$$d = \sum_{i=1}^{4}\left(d_i - \sqrt{(x_i - x)^2 + (y_i - y)^2}\right),$$

where

$$A_1 = d_{12}x_{31} - d_{13}x_{21}; \quad A_2 = d_{12}x_{41} - d_{14}x_{21};$$
$$B_1 = d_{12}y_{31} - d_{13}y_{21}; \quad B_2 = d_{12}y_{41} - d_{14}y_{21},$$
$$C_1 = \frac{1}{2}[d_{13}(d_{12}^2 + x_1^2 - x_2^2 + y_1^2 - y_2^2) - d_{12}(d_{13}^2 + x_1^2 - x_3^2 + y_1^2 - y_3^2)],$$
$$C_2 = \frac{1}{2}\left[d_{14}(d_{12}^2 + x_1^2 - x_2^2 + y_1^2 - y_2^2) - d_{12}(d_{14}^2 + x_1^2 - x_4^2 + y_1^2 - y_4^2)\right].$$

Further elaborations of TDOA are described in [1,7,18]. When the data are in exact, the original system of four squared equations will not, in general, have a real solution such as the TDOA solution.

Appendix B. Source localization in three spatial dimensions

This desideratum is very similar to that treated above, but here one seeks $\mathbf{r} = (x, y, z)$ and d, and one has a spherical-wave source. The global optimum of the respective Δ may occur with two, three or four sgn_s's zeroed. For two sgn_s's zeroed, an analysis analogous to that of Sections 2.1 and 2.3 suffices. For four sgn_s's zeroed, by taking differences, one may eliminate three variables, generating three linear equations. A quadratic equation in the remaining variable yields two candidate solutions, as described in Section 2.2. As above, the cases with more than four sgn_s's zeroed may safely be neglected. Thus, only the case of zeroing three sgn_s's remains, in order to effect the global minimization of Δ.

B.1. Three sgn_s's zeroed (3-d)

Here, using Lagrange multipliers λ, μ and v, we seek the stationary points of

$$\begin{aligned}\tilde{\Delta} = \Delta &+ \lambda((x - x_\ell)^2 + (y - y_\ell)^2 + (z - z_\ell)^2 - (d - d_\ell)^2) + \mu((x - x_m)^2 \\ &+ (y - y_m)^2 + (z - z_m)^2 - (d - d_m)^2) + v((x - x_n)^2 + (y - y_n)^2 \\ &+ (z - z_n)^2 - (d - d_n)^2). \end{aligned} \tag{15}$$

The indices ℓ, m and n are distinct, and, to exhaust this class of stationary points, $\{\ell, m, n\}$ must range over all unordered triples of indices from $\{1, 2, \ldots, S\}$. Thus, given none of the remaining sgn_s's vanish, we seek the stationary points as the solutions of the system:

$$\frac{\partial \tilde{\Delta}}{\partial x} = 2\sum_{s=1}^{S}(x - x_s)w_s\text{sgn}_s + 2\lambda(x - x_\ell) + 2\mu(x - x_m) + 2v(x - x_n) = 0;$$
$$\frac{\partial \tilde{\Delta}}{\partial y} = 2\sum_{s=1}^{S}(y - y_s)w_s\text{sgn}_s + 2\lambda(y - y_\ell) + 2\mu(y - y_m) + 2v(y - y_n) = 0;$$
$$\frac{\partial \tilde{\Delta}}{\partial z} = 2\sum_{s=1}^{S}(z - z_s)w_s\text{sgn}_s + 2\lambda(z - z_\ell) + 2\mu(z - z_m) + 2v(z - z_n) = 0;$$
$$\frac{\partial \tilde{\Delta}}{\partial d} = -2\sum_{s=1}^{S}(d - d_s)w_s\text{sgn}_s - 2\lambda(d - d_\ell) - 2\mu(d - d_m) - 2v(d - d_n) = 0.$$

This system comprises one linear equation in each unknown: x, y, z and d. λ, μ and v will subsequently be selected to effect the three constraints $\text{sgn}_\ell = \text{sgn}_m = \text{sgn}_n = 0$. As above, denote $\sum_{s=1}^{S} w_s\text{sgn}_s$ by ξ, $\sum_{s=1}^{S} x_s w_s\text{sgn}_s$ by ξ_x, $\sum_{s=1}^{S} y_s w_s\text{sgn}_s$ by ξ_y, $\sum_{s=1}^{S} z_s w_s\text{sgn}_s$ by ξ_z and $\sum_{s=1}^{S} d_s w_s\text{sgn}_s$ by ξ_d. Then the foregoing four equations yield, respectively,

$$x = \frac{\xi_x + \lambda x_\ell + \mu x_m + v x_n}{\xi + \lambda + \mu + v}, \quad y = \frac{\xi_y + \lambda y_\ell + \mu y_m + v y_n}{\xi + \lambda + \mu + v},$$
$$z = \frac{\xi_z + \lambda z_\ell + \mu z_m + v z_n}{\xi + \lambda + \mu + v} \quad \text{and } d = \frac{\xi_d + \lambda d_\ell + \mu d_m + v d_n}{\xi + \lambda + \mu + v}.$$

Furthermore, the three constraints yield three quadratic forms: each in a pair of variables. These are unlikely to be definite, and the theory for such systems is rudimentary. Hence, one might proceed by using the quadratic equation to eliminate two of the three variables, using the two equations containing, say, λ to eliminate the other two variables, and seeking the real roots of the remaining equation: an algebraic function of λ (cf. [12]).

References

[1] Aronson EA. Location errors in time of arrival (TOA) and time difference of arrival systems. Sandia Laboratory Technical Report # 77-0495; 1977.

[2] Au WW, Benoit-Bird KJ. Automatic gain control in the echolocation system of dolphins. Nature 2003;423:815.

[3] Bucher R, Misra D. A synthesizable low power VHDL model of the exact solution of three dimensional hyperbolic positioning system. VLSI Des 2002;15:507–10.

[4] Gauss CF. Theoria combinationis observationum erroribus minimis obnoxiae [Translated by G.W. Stewart]. Philadelphia: SIAM Press; 1995.

[5] Harley HE, Putman EA, Roitblat HL. Bottlenose dolphins perceive object features through echolocation. Nature 2003;424:667–9.

[6] Hazewinkel M. Encyclopædia of mathematics. Dordrecht: Kluwer; 1991.

[7] Ho KC, Chan YT. Solution and performance analysis of geolocation by TDOA. IEEE Trans Aerosp Electron Syst 1993;29:1311–22.
[8] Huang Y, Benesty J, Elko GW. An efficient linear-correction least-squares approach to source localization. In: Proceedings of the 2001 IEEE workshop on the applications of signal processing to audio and acoustics; 2001. p. 21–4.
[9] Hurt NE. Maximum likelihood estimation and MUSIC in array localization signal processing: a review. Multidimens Syst Signal Process 1990;1:279–325.
[10] Landau BV, West M. Estimation of the source location and the determination of the 50% probability zone for an acoustic source locating system (SLS) using multiple systems of 3 sensors. Appl Acoust 1997;52:85–100.
[11] Lanman DR. Distributed sensor networks with collective computation. Los Alamos National Laboratory Technical Report #01-4388; 2001.
[12] Mora T. Solving polynomial equation systems I. In: The encyclopedia of mathematics and its applications. Cambridge: Cambridge University Press; 2003.
[13] Nemzek RJ, Dreicer JS, Torney DC, Warnock TT. Distributed sensor networks for detection of mobile radioactive sources. IEEE Trans. Nucl. Sci. 2004;51:1693–700.
[14] Peremans H, Walker VA, Papadopoulos G, Hallam JCT. Evolving batlike pinnae for target localization by an echolocator. Lect Notes Comput Sci 1998;1478:230–9.
[15] Pfeiffer EW, Johnstone B, Landau S. Where ware. Technol Rev 2003;106(7):46–52.
[16] Pinch ER. Optimal control and the calculus of variations. Oxford: Oxford University Press; 1993.
[17] Press WH. Numerical recipes in Fortran 77. 2nd ed. Cambridge: Cambridge University Press; 1992. 698.
[18] Schmidt R. Least squares range difference location. IEEE Trans Aerosp Electron Syst 1996;32:234–42.
[19] Suga N. Biosonar and Neural Computation in Bats, Scientific American 1990;262:60–8.